# How to Get More Memory for your PBS Job

UPDATE IN PROGRESS: Aitken was recently expanded with 1,024 AMD Rome nodes. This article is being updated to support use of the new nodes.
If your job terminates because the nodes it is running on do not have enough available memory, try using one or more of the methods described below.

Note: The memory listed for each type of node in this article refers to the total physical memory in the node. The amount of memory that is actually available to a PBS job is slightly less than the total physical memory of the nodes it runs on, because the operating system can use up to 4 GB of memory in each node. At the beginning of a job, the PBS prologue checks the nodes for free memory and guarantees that 85% of the total physical memory is available for the job.

## Use A Processor With More Memory

The following list shows the amount of total physical memory per node for the various Pleiades, Aitken, and Electra processor types:

Sandy Bridge
: 32 GB

Ivy Bridge
: 64 GB

Haswell
: 128 GB

Broadwell
: 128 GB

Skylake
: 192 GB

Cascade Lake
: 192 GB

If your job runs of out memory on a Sandy Bridge node, you can try using one of the other node types instead. For example, change:

```
#PBS -lselect=1:ncpus=16:model=san
```

to one of the following:

```
#PBS -lselect=1:ncpus=20:model=ivy

#PBS -lselect=1:ncpus=24:model=has

#PBS -lselect=1:ncpus=28:model=bro

#PBS -lselect=1:ncpus=40:model=sky_ele

#PBS -lselect=1:ncpus=40:model=cas_ait
```

## Use LDANs as Sandy Bridge Bigmem Nodes

The Lou data analysis nodes (LDANs) can be dynamically switched between two modes: *LDAN mode* and *bigmem mode*. A PBS job on ldan[*11-12*] can access 768 GB of memory per node; on ldan[*13-14],* a PBS job can access 1.5 TB per node.

## LDAN mode

In LDAN mode, the home filesystem of an LDAN assigned to your job is set to your Lou home filesystem in order to facilitate processing of data on Lou.

For LDAN mode, specify **`-q ldan`**.

## Bigmem mode

In bigmem mode, the home filesystem of an LDAN assigned to your job is set to your Pleiades home filesystem. This allows the LDANs to be used as Sandy Bridge bigmem nodes.

To use bigmem mode, specify one of the standard PBS queues (such as **`devel`**, **`debug`**, **`normal`**, or **`long`**) together with **`:model=ldan`**. For example:

```
#PBS -q normal
#PBS -lselect=1:ncpus=16:model=ldan+10:ncpus=20:model=ivy

#PBS -q normal
#PBS -lselect=1:ncpus=16:model=ldan:mem=250GB+10:ncpus=16:model=san
```

## Use Ivy Bridge Bigmem Nodes

Three Ivy Bridge nodes have 128 GB/node of memory instead of the standard 64 GB/node. To request these nodes, specify **`bigmem=true`** and **`model=ivy`**. For example:

```
#PBS -lselect=1:ncpus=20:model=ivy:bigmem=true
```

Note: The Ivy Bridge bigmem nodes are not available for jobs in the **`devel`** queue.

## Run Fewer Processes per Node

If all processes use approximately the same amount of memory, and you cannot fit 16 processes per node (Sandy Bridge); 20 processes per node (Ivy Bridge); 24 processes per node (Haswell), 28 processes per node (Broadwell), or 40 processes per node (Skylake/Cascade Lake), then reduce the number of processes per node and request more nodes for your job.

For example, change:

```
#PBS -lselect=3:ncpus=20:mpiprocs=20:model=ivy
```

to the following:

```
#PBS -lselect=6:ncpus=10:mpiprocs=10:model=ivy
```

Recommendation: Add the command **`mbind.x -gm -cs`** before your executable in order to (1) monitor the memory usage of your processes and (2) balance the workload by spreading the processes between the two sockets of each node. For example:

```
mpiexec -np 60 /u/scicon/tools/bin/mbind.x -gm -cs ./a.out/
```

If you are running a typical MPI job, where the rank 0 process performs the I/O and uses a large amount of buffer cache, assign rank 0 to one node by itself. For example, if rank 0 requires up to

20 GB of memory by itself, change:

```
#PBS -lselect=1:ncpus=16:mpiprocs=16:model=san
```

to the following:

```
#PBS -lselect=1:ncpus=1:mpiprocs=1:model=san+1:ncpus=15:mpiprocs=15:model=san
```

If the rank 0 process needs 20-44 GB of memory by itself, use:

```
#PBS -lselect=1:ncpus=1:mpiprocs=1:bigmem=true:model=ivy+1:ncpus=19:mpiprocs=19:model=ivy
```

#### Use the Endeavour System

If any process or thread in a multi-process or multi-thread job needs more than about 250 GB, the job won't run on Pleiades. Instead, run it on Endeavour, which is a shared-memory system.

## Report Bad Nodes

If you suspect that certain nodes have less total physical memory than normal, report them to the NAS Control Room staff at (800) 331-8737, (650) 604-4444, or by email at support@nas.nasa.gov. The nodes can be taken offline while we resolve any issues, preventing them from being used before they are fixed.

-->

---

Article ID: 222
Last updated: 12 Mar, 2021
Revision: 69
Optimizing/Troubleshooting -> PBS Jobs -> How to Get More Memory for your PBS Job
https://www.nas.nasa.gov/hecc/support/kb/entry/222/